Morphologically-conditioned tonotactics in multilevel Maximum Entropy grammar

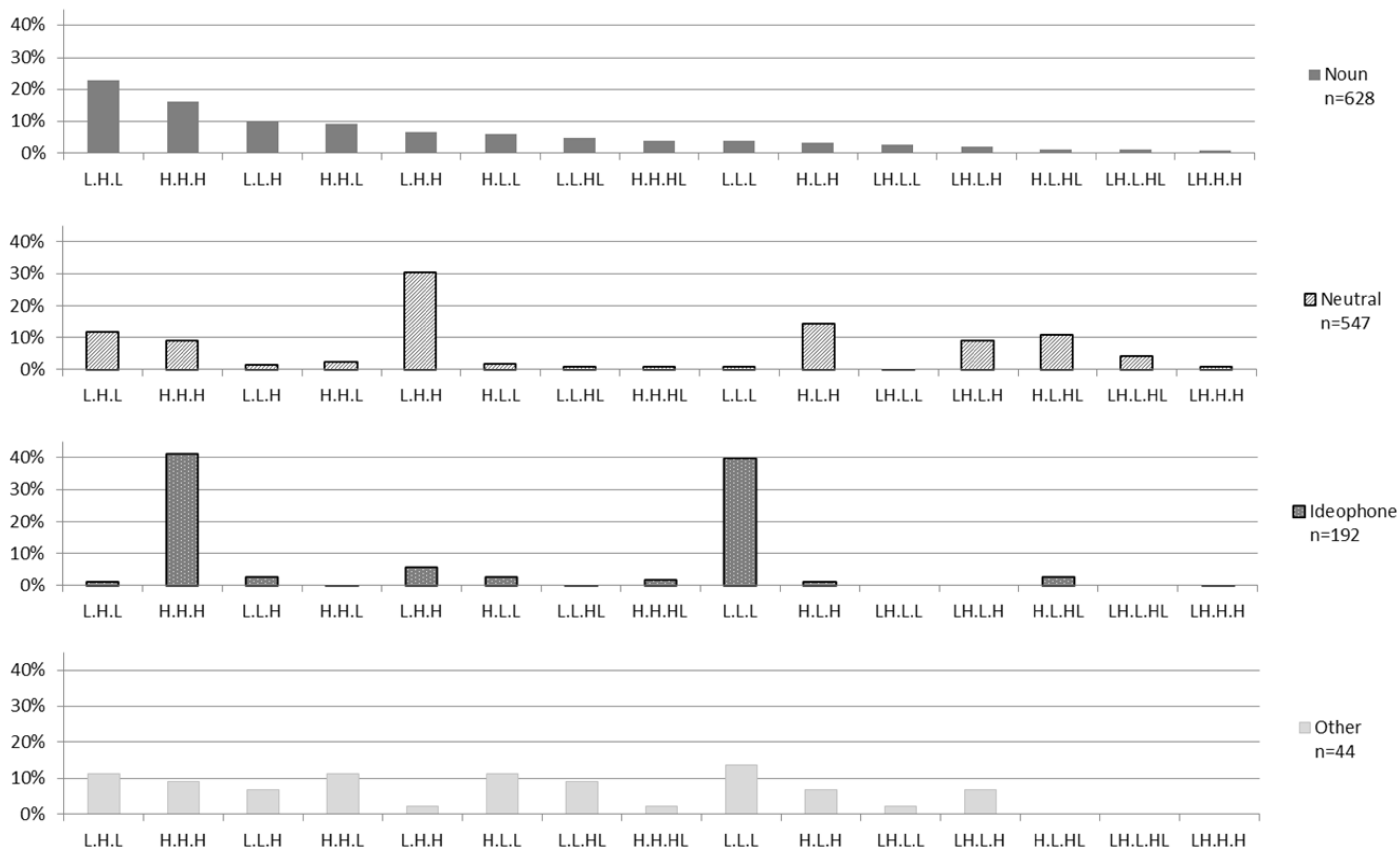
Stephanie S Shih
University of California, Merced
shih@ucmerced.edu

Sharon Inkelas
University of California, Berkeley
inkelas@berkeley.edu

1 INTRODUCTION

❖ This talk presents a case of lexically-conditioned tonotactics variation from Mende (Mande, Sierra Leone).

(1) Top trisyllabic surface tone patterns in the Mende lexicon



- Part of speech-sensitive patterns =
  - different lexical classes can exhibit different phonological patterns.
  - noted cross-linguistically: e.g., English noun versus verb stress patterns
  - applications in e.g., comprehension, parsing

- Two hypotheses about how lexical class-conditioned sensitivity could work:

  A. Lexical class differences are limited by the grammar/UG.
     - differences only in faithfulness, not markedness (e.g., Ito & Mester 1995; Alderete 2001; Smith 2011)
     - preferential classes, e.g., nouns // verbs will show a subset of noun patterns (Smith 2011)

  B. Lexical classes can each have their own completely independent phonological profiles (e.g., Ito & Mester 1995; Inkelas & Zoll 2007; Anttila 2002; Pater 2009).
     ↳ We argue for this latter hypothesis here.

❖ This talk:
- quantitatively models space of lexically-conditioned variation and frequency of variation across the corpus,
- using a ‘varying slopes’ approach in Maximum Entropy Harmonic Grammar (MaxEnt; Goldwater & Johnson 2003; Wilson 2006; Jäger 2007; Hayes & Wilson 2008; a.o.).
- ↳ This approach directly addresses the overarching problem in morphophonology of how to quantify the heterogeneity that morphological conditioning can engender in a phonological system.

2 MENDE TONOTACTICS

2.1 History, early generative accounts

- Early generative accounts of Mende noticed common tone patterns recurrent in the language, particularly in nouns (Leben 1978) → ‘tone melodies’ (see also Hyman 1987 for similar tone melodies in Kukuya),
- In Autosegmental Phonology, these surface tone patterns were modeled using
  - geometric association conventions of Autosegmental Phonology
  - 5 underlying tone melodies (H, L, HL, LH, LHL) as source of all surface patterns.

(2) ndàvúlá ‘sling’ *L→R, 1↔1 association, then spread.*
L H

- But, subsequent work pointed out many surface patterns that deviate from the supposed five melodies or their “universal” autosegmental association principles (Dwyer 1978; Conteh et al. 1983; Zoll 2003; Zhang 2007).

(3) a. lèlèmá ‘praying mantis’ *violates association principles*
L H

b. gbágbɛ̌mà ‘sensitive plant’ *cannot arise from one of the 5 tone melodies; violates association principles*
H LH L

2.2 Data

- Mende dictionary: *n*=5,412 (Innes 1969)
- 1 to 3-syllable words: *n*=4,989
  - Morpheme breaks are not indicated in Innes, but a primary source of morphological complexity (in nouns, at least) appears to be total reduplication in 4-syllable words, which we’re not looking at here.
- Parts of speech
  - Nouns 2,494
  - Neutrals 1,442 (verbs, adjectives)
  - Ideophones 762
  - Other 291 (pronouns, conjunctions, interjections, adverbs, etc.)

2.3 A fresh look at tone: the theoretical underpinnings

- Agreement by Correspondence Theory (ABC; Hansson 2001; Rose & Walker 2004; Bennett 2013; a.o.)
  - grounded in basic principles of similarity and proximity attraction, modeling instability in syntagmatic phonological relationships (Wayment 2009; Inkelas & Shih 2013).
  - Elements that are sufficiently similar/proximal interact in e.g., assimilation/dissimilation.
  - For simplicity in this talk, constraints are reformulated into more familiar phonotactic markedness format (but it’s still ABC under the hood; cf. Hansson 2014).

- Q Theory (e.g., Shih & Inkelas 2014)
  - decomposes segments into strings of (2 or 3) smaller, temporally-sequenced, featurally-uniform subsegments, which bear tone features.

(4) a. Q → (q^1 q^2)
b. ă → (à á)
c. $\widehat{\text{LH}}$ → (L H)

– provides a more fine-grained point of reference for the grammar: crucial for e.g., contour tones.
– divorces issue of what are the minimal units that carry tone features versus what are the units that participate in tonal alternations/phenomena.

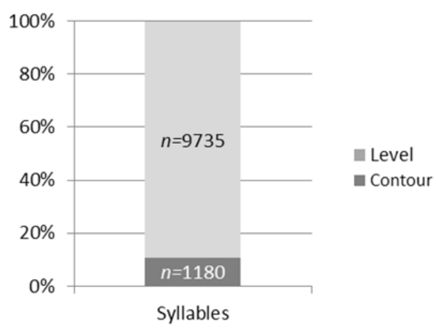
- Basic relevant differences between Autosegmentalism (e.g., Leben 1973) and ABC+Q for this talk:
  – Constraints grounded in principles of similarity- and proximity-based interaction.
  – No reliance on geometric, autosegmental 'line' representations
  – No reliance on operations that reference autosegmental 'lines': i.e., tone association rules.[1]

2.4 Observed patterns for Mende surface tones

- Primary observations taken from Inkelas & Shih 2015.

2.4.1 Contour toned syllables (and tone transitions in general—except at syllable boundaries) are avoided.

(5)



- Significantly fewer contour syllables than expected, if syllable tone patterns could be HH, LL, LH, and HL. χ^2=6705.270, p<0.0001.

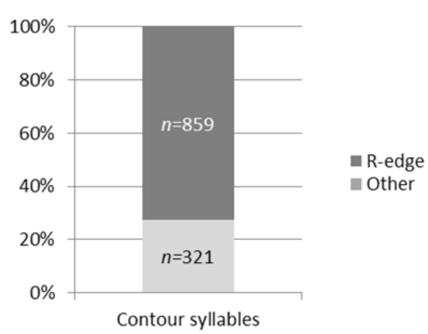
(6) *[αT]::[βT] (*CHANGE) Penalise every sequence of adjacent *q*'s that are tonally non-identical.

(7)

σ.σ	*freq*	*CHANGE *[αT]::[βT]
☞ a. LL.HH	701	1
b. LL.HL	389	W2

2.4.2 If necessary, contour tones are tolerated at the right edge.

(8)



- The significant majority of contour syllables occur at the right edge of words.
- versus equal probability across all syllables: χ^2=353.407, p<0.0001.

[1] This possibility was hinted at by Zoll (2003), but not fully explored.

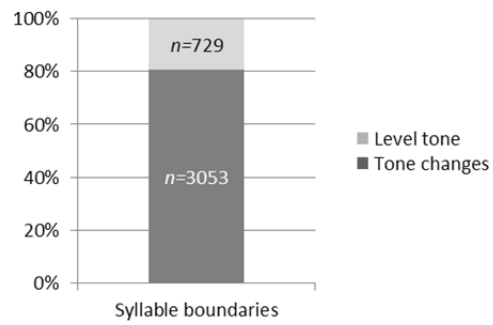
(9) $*([\alpha T]::[\beta T])_{\sigma w}$ (*WeakCONTOUR) Penalise every adjacent *q* sequence of non-identical tones within a weak (i.e., non-final) syllable. (cf. COINCIDE-CONTOUR; Zoll 2003)

(10)

	σ.σ	*freq*	*WeakCONTOUR $*([\alpha T]::[\beta T])_{\sigma w}$	*CHANGE $*[\alpha T]::[\beta T]$
☞ a.	LL.HL	389		2
b.	LH.LL	77	W1	2

2.4.3 Tone changes align with syllable boundaries

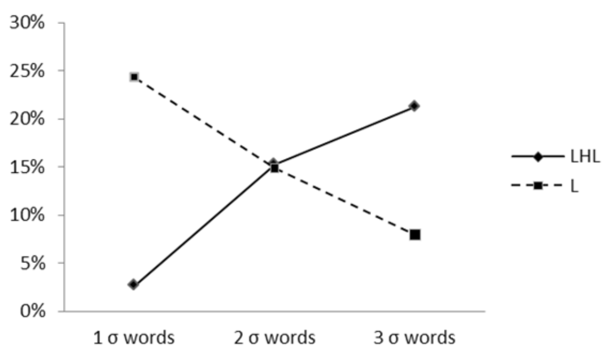
(11)



- Syllable boundaries coincide significantly with tone changes (in polysyllabic words that have non-level surface patterns) . χ^2=1428.074, p<0.0001

- Resulting prediction: tone melody complexity should correlate with the number of syllables in a word. More syllables → more non-level tone patterns (Inkelas & Shih 2015).

(12)



- As # of syllables in a word increases, so does the percentage of more complex melodies. Conversely, less complex, all-level melodies decrease.

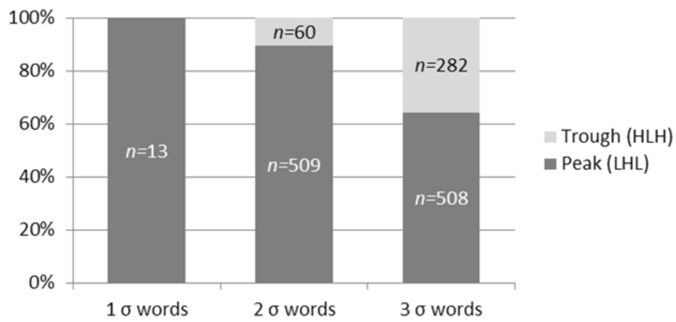
(13) $*[\alpha T]:\$:[\alpha T]$ (CHANGE@$) Penalise every adjacent, tonally-identical *q* sequence that is separated by a syllable boundary ($). (cf. xx-Edge constraints; Bennett 2013)

(14)

	σ.σ.σ	*freq*	CHANGE@$ $*[\alpha T]:\$:[\alpha T]$	*WeakCONTOUR $*([\alpha T]::[\beta T])_{\sigma w}$	*CHANGE $*[\alpha T]::[\beta T]$
☞ a.	LL.HH.LL	233			2
b.	LL.HH.HH	218	W1		L1
c.	LL.LL.LL	112	W2		L
d.	LL.LL.HH	78	W1		L1
e.	LH.LL.HH	64		W1	W3
f.	LH.HH.LL	7	W1	W1	2

2.4.4 HLH troughs are avoided

(15)

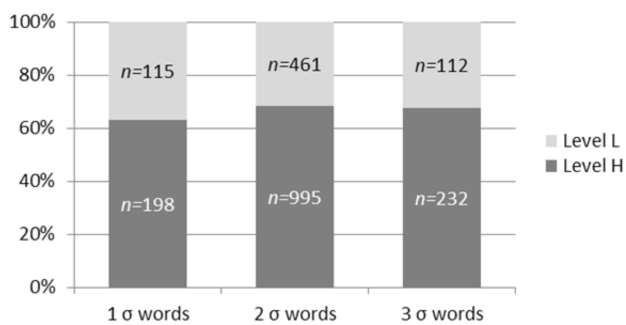


- Troughs are underrepresented, but their presence increases as the number of syllables in words increases.

(16) *TROUGH Penalise any two H tones separated by any number of consecutive L tones. (Cahill 2007; a.o.)

2.4.5 Words preferably have at least one H tone

(17)



- Level H tone patterns (e.g., HH.HH.HH) are more common than Level L tone patterns (e.g., LL.LL.LL).

(18) HAVEH Penalise any word with no H tone.

2.5 Constraint summary

(19)

Constraint	Nickname	Definition
*[αT]::[βT]	*CHANGE	Penalise every sequence of adjacent *q*’s that are tonally non-identical.
$*([\alpha T]::[\beta T])_{\sigma w}$	*WeakCONTOUR	Penalise every adjacent *q* sequence of non-identical tones within a weak (i.e., non-final) syllable.
*[αT]:$:[αT]	CHANGE@$	Penalise every adjacent, identical *q* sequence that is separated by a syllable boundary.
*TROUGH	*TROUGH	Penalise any two H tones separated by any number of consecutive L tones.
HAVEH	HAVEH	Penalise any word with no H tone.

- Note that no matter the OT analysis, the goal in this paper is to examine and model the variance of part-of-speech-specific tonotactics. (Almost) any constraint set would be fine for this, and the fact that there’s lexically-conditioned differences (see (1)) doesn’t change.

3 ANALYSIS

3.1 Maximum Entropy Harmonic Grammar

- Maximum Entropy Harmonic Grammar (MaxEnt) (Goldwater & Johnson 2003; Wilson 2006; Jäger 2007; Hayes & Wilson 2008; a.o.), fitted using MaxEnt Grammar Tool (Hayes et al. 2009).[2]
  - ranks probabilities (i.e., comparative grammaticality) of outcome candidates in variable data.
- Output = probability distribution over all possible surface tone pattern combinations of LL, HH, LH, HL syllables (LHL, HLH are found on monosyllables only)

3.2 Cophonologies/Indexed Constraints as Varying Slopes

- Varying slopes = additive weight adjustment for constraints in the Base Grammar, per each lexical class.
  - Such an approach has been hinted at before: e.g., Albright 2008; Coetzee & Pater 2011.

(20) $$w_1C_1 + w_2(C_1 \times \textsc{Noun}) + w_3(C_1 \times \textsc{Neut}) + w_4(C_1 \times \textsc{Id}) + \cdots + w_kC_{iN}$$ (cf. fn 2)

Base grammar	=	w_1C_1
Adjusted noun grammar	=	$w_1C_1 + w_2(C_1 \times \textsc{Noun})$
Adjusted neutral grammar	=	$w_1C_1 + w_3(C_1 \times \textsc{Neut})$
Adjusted ideophone grammar	=	$w_1C_1 + w_4(C_1 \times \textsc{Id})$

- Base Grammar predicts overall tonotactics for the language, and lexical class-specific tonotactics are predicted by the additive combination of weights of Base Grammar constraints and class-specific weights.
- Cophonologies ≅ bundles of Indexed Constraints.
- Because there is a finite set of lexical classes, varying slopes are formally executed here as interaction terms.[3]
  - for other previous uses of interaction terms in MaxEnt HG, usually as weighted constraint conjunction: see e.g., Hayes et al. 2012; Pater & Moreton 2012; Green & Davis 2014; Shih, to appear.

3.3 What is the base grammar?

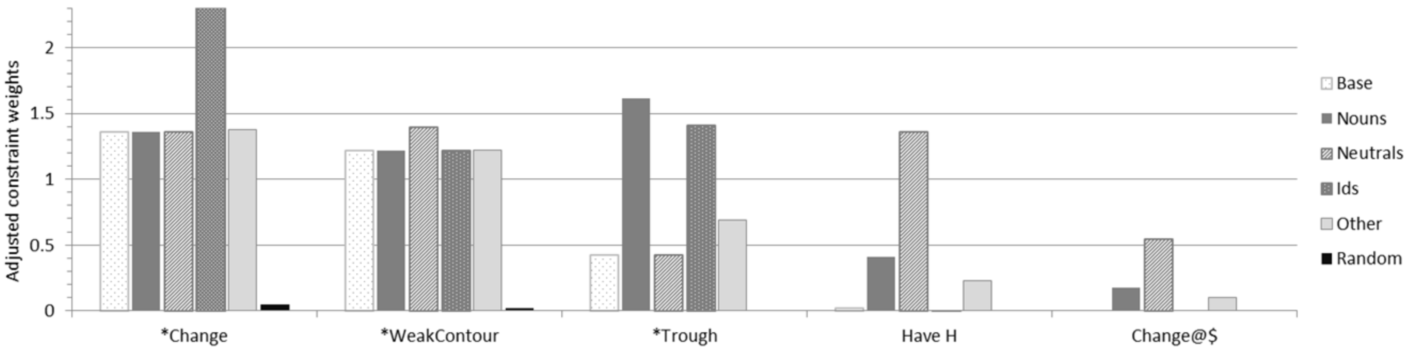
- Trained a base grammar on bootstrapped samples (*n*=2500 / sample) of possible tone melodies in Mende, drawn from the existing distribution.
  - mirrors a Mende speaker knowing word tone patterns but not (yet) part of speech.
  - did not want to *a priori* assume one lexical class as "default."
  - across base grammar samples, the results are stable; a representative result is reported here.

- A grammar trained without any data for base grammar provides very similar results (not reported here).
- Another baseline: grammar on a totally random tone patterns (*n*=5000), generated from random combinations of HH, LL, LH, and HL syllables for 1–3 syllable words is also provided here for comparison.

[2] $\mathcal{H}\text{armony}(x, i) = \sum_{k=1}^{N} w_k C_i(x) = \boldsymbol{w_k} \cdot \boldsymbol{C_i}$; where x = candidate for input i, w_k= weight of constraint C_i, $C_i(x)$ = number of violations of C_i that x incurs, and N = vector of constraints $(C_{i1} \ldots C_{iN})$.
[3] For justification on the equivalency of interaction terms and random slopes in this situation, see e.g., Gelman & Hill 2007.

4 RESULTS

(21) Adjusted MaxEnt weights, in graphical form
(adjusted wC_i = base wC_i + indexed wC_i)



(22) Adjusted MaxEnt weights, in numbers

	Base	Adjusted Grammars				Random
		Nouns	*Neutrals*	*Ideophones*	*Other*	
*CHANGE	1.356	1.356	1.356	2.537	1.374	0.047
*WeakCONTOUR	1.220	1.220	1.391	1.220	1.220	0.018
*TROUGH	0.426	1.612	0.426	1.409	0.691	0
HAVE H	0.017	0.411	1.361	0.017	0.232	0
CHANGE@$	0	0.175	0.543	0	0.099	0

(23) Constraint weights by lexical class (top→bottom: highest weight→lowest weight)

Base	Cophonology Indexed Grammars			
	Nouns	*Neutrals*	*Ideophones*	*Other*
*CHANGE	*TROUGH	*WeakCONTOUR	*CHANGE	*CHANGE
*WeakCONTOUR	*CHANGE	HAVE H	*TROUGH	*WeakCONTOUR
*TROUGH	*WeakCONTOUR	*CHANGE	*WeakCONTOUR	*TROUGH
HAVE H	HAVE H	CHANGE@$	CHANGE@$	HAVE H
CHANGE@$	CHANGE@$	*TROUGH	HAVE H	CHANGE@$

4.1 Results and observations

4.1.1 Sanity check

- A MaxEnt grammar on completely random data yields constraint weights pretty close to 0.

4.1.2 Base grammar

- Reveals importance of general contour tone avoidance and R-edge contour tone alignment.
  - *CHANGE = 1.356
  - *WeakCONTOUR = 1.220
  - cf. "Random" model results, where weights are much lower, nearly at 0.
- These two constraints remain highly weighted across all lexical classes.
  - weight of *WeakCONTOUR stays stable across the lexicon, suggesting that it's a very general principle.
- Reflects universal dispreferences for (nonfinal) contour tones (e.g., Gordon 2001; Zhang 2004).

4.1.3 Syllable edge – Tone change alignment

- Only nouns and neutrals show an effect of syllable edge-tone transition alignment. → see adjusted weights for CHANGE@$:
  - CHANGE@$×NOUN = 0.175
  - CHANGE@$×NEUT = 0.543
  - versus CHANGE@$×ID = 0

- Shows that ideophones often involve much simpler surface tone patterns than the other lexical classes.
  - e.g., *vòvòlò* ‘creaking’
  - In addition, ideophones significantly increase the weight of *CHANGE

(24)

σ.σ.σ, ID	*freq*	*CHANGE×ID *[αT]::[βT]×ID w_{adj}=2.537	CHANGE@$×ID *[αT]:$:[αT]×ID w_{adj}=0	$\mathcal{H}$
a. LL.HH.LL	2	W2		**5.073**
☞ b. LL.LL.LL	76		L2	**0**

4.1.4 HAVE H*: a verb thing*

- Neutrals in particular (but also nouns to a certain extent) show a much greater affinity for a requisite H tone than ideophones.
  - = level L surface patterns much more tolerated for Ideophones > Nouns > Neutrals.

(25)

σ.σ.σ, NEUT	*freq*	HAVEH×NEUT w_{adj}=1.360656	$\mathcal{H}$
a. LL.LL.LL	5	1	**1.36**

σ.σ.σ, NOUN		HAVEH×NOUN w_{adj}=0.410611	
b. LL.LL.LL	25	1	**0.41**

σ.σ.σ, ID		HAVEH×ID w_{adj}=0.016843	
c. LL.LL.LL	76	1	**.02**

4.1.5 *TROUGH*: not a verb thing*

- Neutrals also differ from Nouns and Ideophones by showing greater preference for tone transitions that align with syllable boundaries, even if these result in a HLH trough.
  - e.g., *húvùndí* ‘be mouldy, mildewed’
  - = Greater dispreference of troughs in Nouns and Ideophones.
  - Much lower adjusted weight for *TROUGH in Neutrals than other lexical classes:
  - *TROUGH ×NEUT = 0.426
  - *TROUGH ×NOUN = 1.612
  - *TROUGH ×ID = 1.409

- Perhaps a contrast maximization issue?

(26)

		σ.σ.σ, Neut	*freq*	*Trough×Neut w_{adj}=0.426	Change@$ *[αT]:$:[αT]×Neut w_{adj}=0.543	$\mathcal{H}$
a.	☞ i.	HH.LL.HH	78	1		**0.43**
	ii.	HH.HH.HH	48		**2**	**1.09**

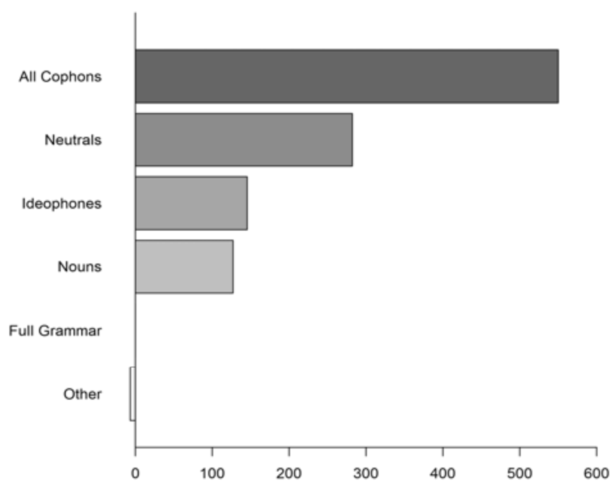
		σ.σ.σ, Noun		*Trough×Noun w_{adj}=1.612	*[αT]:$:[αT]×Noun w_{adj}=0.175	
b.	i.	HH.LL.HH	21	1		**1.61**
	☞ ii.	HH.HH.HH	101		**2**	**0.35**

		σ.σ.σ, Id		*Trough×Id w_{adj}=1.409	*[αT]:$:[αT]×Id w_{adj}=0	
c.	i.	HH.LL.HH	2	1		**1.41**
	☞ ii.	HH.HH.HH	79		**2**	**0.00**

4.2 Importance of lexical conditioning

- Are the cophonology/indexed grammars contributing significantly to improving our base grammar?
- Tested here using Akaike Information Criterion (AIC_C) model comparison.[4]
  - How much information is lost when cophonology/indexed constraints are removed?

(27) ΔAIC_C from full model for each cophonology removed
= how much information is lost for each cophonology is in turn removed.[5]



- AIC_C comparison results show that lexical class-specific cophonologies/indexed constraints produce a better fit for the data.
  - Exception that is expected: the “Other” part of speech class, which is heterogeneous as is.
- Certain lexical classes provide more information on top of the baseline, suggesting ‘distance’ from the full model.
  - Neutral-specific cophonology contributes a lot of additional information to the grammar.
  - Interestingly, ideophones are *not* the most extreme lexical class.
- For future work: comparing individual indexed constraints versus entire cophonologies of indexed constraints.

[4] Second-order AIC_C is used here, because it adjusts more strictly for differences in the number of constraints between comparison grammars. See e.g., Burnham & Anderson 2002 for justification of AIC_C-based model comparison. For use in comparing phonological grammars, see e.g., Shih, to appear (cf. Wilson & Obdeyn 2009).

[5] $\Delta AIC_C \geq 10$ is considered large (equivalent to 150:1 that the second best model has essentially no support of being as good as the best model).

5 IMPLICATIONS & CONCLUSION

Lexically-conditioned phonology

- Part-of-speech phonological differences go beyond noun-adjective-verb distinctions (as observed by e.g., Smith 2011).
  - ↳ more closely resemble the complexity of morphophonological alternations and variation.

- Lexically-conditioned phonology is not just a matter of differential faithfulness (e.g., $\text{FAITH}_{\text{NOUN}} \gg \text{FAITH}_{\text{VERB}}$; Smith 2011).
- Overall, indexation of markedness constraints per lexical class gains better fit for the data.
- There are hints of markedness reversals of the kind that 'Grammar Dependence' (Alderete 2001) predicts impossible (cf. Pater 2009):
  - – E.g., Across nearly all of the grammars (base and lexical class-specific) in Mende, *TROUGH is ranked fairly highly w/r/t the other constraints.
  - – In Neutrals, however, *TROUGH is weighted at the bottom.
  - ↳ Therefore, a structure (e.g., HLH) that might otherwise be highly marked in the rest of the language is quite good and (comparatively) unmarked for Neutrals.

Quantifying morphological conditioning

- Our approach can quantify how much lexical class-specific phonotactics can differ from the rest of the lexicon.
  - – E.g., how different can ideophone phonology be? → long standing issue for e.g., African languages (see Rose 2015 for recent summary).
  - – Results show that ideophones operate within fairly conservative parameters of the overall Mende tonotactics grammar—cf., neutrals.
  - – Potentially useful as a scale of ideophone idiosyncrasy (cf. Newman 2001 on Hausa ideophones).
- Because we *a priori* restrict standard HG/OT grammars to positive weights (i.e., to only penalise rather than reward structures), we're not really examining the full potential space of variation between cophonologies yet.
- In the future, this 'varying slopes' approach could extend to other types of morphology (e.g., affixation; dominant tone melodies): e.g., what are the tone patterns that are grammaticalised in a language as productive morphemes?

For the future

- What is the utility of these phonological differences between groups?
  - – There's evidence from other languages that speakers are capable of learning these lexical statistics (e.g., Coetzee 2014),
  - – and that such phonological differences help in e.g., acquisition, processing, information load (e.g., Monaghan et al. 2010).

Acknowledgements to Laura McPherson, Brian Smith, Alan Yu, Kie Zuraw, and audiences at UC Berkeley, UC San Diego, and CLS 2015 for discussion on various portions of this work.

To save trees, references are available online.
http://cogsci.ucmerced.edu/shih/ShihInkelas_AMP2015handout.pdf

6 REFERENCES

ALBRIGHT, ADAM. 2008. How many grammars am I holding up? Discovering differences between word classes. *Proceedings of the 26th West Coast Conference on Formal Linguistics*, ed. by Charles B. Chang and Hannah J. Haynie, 1–20. Somerville, MA: Cascadilla Proceedings Project.

ALDERETE, JOHN D. 2001. Dominance Effects as Transderivational Anti-Faithfulness. *Phonology* 18.201–253.

ANTTILA, ARTO. 2002. Morphologically Conditioned Phonological Alternations. *Natural Language & Linguistic Theory* 20.1–42.

BENNETT, WILLIAM G. 2013. Dissimilation, consonant harmony, and surface correspondence. New Brunswick, NJ: Rutgers, The State University of New Jersey ph.d. dissertation.

BURNHAM, KENNETH P.; and DAVID R. ANDERSON. 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. 2nd edition. New York, NY: Springer-Verlag.

CAHILL, MIKE. 2007. More universals of tone. Ms. SIL, ms.

COETZEE, ANDRIES W; and JOE PATER. 2011. The place of variation in phonological theory. *The Handbook of Phonological Theory*, ed. by John A Goldsmith, Jason Riggle, and Alan C. L. Yu, 401–434. John Wiley & Sons.

COETZEE, ANDRIES W. 2014. Grammatical change through lexical accumulation: Voicing cooccurrence restrictions in Afrikaans. *Language* 90.693–721.

CONTEH, P.; E. COWPER; D. JAMES; K. RICE; and M. SZAMOSI. 1983. A reanalysis of tone in Mende. *Current Approaches to African Linguistics*, ed. by J. Kaye, H. Koopman, D. Sportiche, A. Dugas, J. Kaye, H. Koopman, D. Sportiche, and A. Dugas, 2:127–137. Dordrecht: Foris.

DWYER, DAVID. 1978. What sort of tone language is Mende? *Studies in African Linguistics* 9.167–208.

GELMAN, ANDREW; and JENNIFER HILL. 2007. *Data analysis using regression and multilevel/hierarchical models*. Analytic methods for social research. New York, NY: Cambridge University Press.

GOLDWATER, SHARON; and MARK JOHNSON. 2003. Learning OT Constraint Rankings Using a Maximum Entropy Model. *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, ed. by Jennifer Spenader, Anders Eriksson, and Östen Dahl, 111–120. Stockholm: Stockholm University.

GORDON, MATTHEW. 2001. A typology of contour tone restrictions. *Studies in Language* 25.405–444.

GREEN, CHRISTOPHER; and STUART DAVIS. 2014. Superadditivity and Limitations on Syllable Complexity in Bambara Words. *Perspectives on Phonological Theory and Acquisition: Papers in Honor of Daniel Dinnsen*, ed. by Ashley W. Farris-Trimble and Jessica Barlow. Philadelphia/Amsterdam: John Benjamins Co.

HANSSON, GUNNAR ÓLAFUR. 2001. Theoretical and Typological Issues in Consonant Harmony. University of California, Berkeley ph.d. dissertation.

HANSSON, GUNNAR ÓLAFUR. 2014. (Dis)Agreement by (Non)Correspondence: Inspecting the foundations. *UC Berkeley Phonology Lab Annual Report: ABC Conference Archive*, 3–62. Berkeley, CA. http://linguistics.berkeley.edu/phonlab/annual_report/documents/2014/annual_report_2014_ABCC.html.

HAYES, BRUCE; and COLIN WILSON. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39.379–440.

HAYES, BRUCE; COLIN WILSON; and BEN GEORGE. 2009. *Maxent Grammar Tool*. http://www.linguistics.ucla.edu/people/hayes/MaxentGrammarTool/.

HAYES, BRUCE; COLIN WILSON; and ANNE SHISKO. 2012. Maxent grammars for the metrics of Shakespeare and Milton. *Language* 88.691–731.

HYMAN, LARRY M. 1987. Prosodic Domains in Kukuya. *Natural Language & Linguistic Theory* 5.311–333.

INKELAS, SHARON; and STEPHANIE S SHIH. 2013. Unstable surface correspondence as the source of local conspiracies. Paper presented at the NELS 2013, University of Connecticut. http://stanford.edu/~stephsus/InkelasShih_NELS_2013_handout.pdf.

INKELAS, SHARON; and STEPHANIE S SHIH. 2015. Tone melodies in the age of Surface Correspondence. Paper. Paper presented at the 51st Annual Meeting of the Chicago Linguistic Society (CLS), University of Chicago.

INNES, GORDON. 1969. *A Mende-English Dictionary*. London and New York: Cambridge University Press.

JÄGER, GERHARD. 2007. Maximum Entropy Models and Stochastic Optimality Theory. *Architectures, Rules, and Preferences: A Festschrift for Joan Bresnan*, ed. by Jane Grimshaw, Joan Maling, Chris Manning, Jane Simpson, and Annie Zaenen. Stanford, CA: CSLI.

KULLBACK, SOLOMAN; and RICHARD A. LEIBLER. 1951. On Information and Sufficiency. *Annals of Mathematical Studies* 2.79–86.

LEBEN, WILL. 1978. The representation of tone. *Tone: A Linguistic Survey*, ed. by Victoria Fromkin, 177–219. New York: Academic Press.

LEBEN, WILLIAM R. 1973. Suprasegmental phonology. Cambridge, MA: Massachusetts Institute of Technology ph.d. dissertation.

MONAGHAN, PADRAIC; MORTEN H. CHRISTIANSEN; THOMAS A. FARMER; and STANKA A. FITNEVA. 2010. Measures of phonological typicality: Robust coherence and psychological validity. *The Mental Lexicon* 5.281–299.

NEWMAN, PAUL. 2001. Are ideophones really as weird and extra-systematic as linguists make them to be? *Ideophones*, ed. by F. K. Erhard Voeltz and Christa Kilian-Hatz, 251–258. Typological Studies in Language 44. Amsterdam: John Benjamins.

PATER, JOE. 2009. Morpheme-specific phonology: constraint indexation and inconsistency resolution. *Phonological argumentation: essays on evidence and motivation*, ed. by Steve Parker. Advances in Optimality Theory. Equinox.

PATER, JOE; and ELLIOT MORETON. 2012. Structurally biased phonology: Complexity in learning and typology. *The English and Foreign Languages Journal* 3.1–44.

ROSE, SHARON. 2015. Phonology of Ideophones in African Languages. Paper. Paper presented at the WOCAL 8, Kyoto University, Japan.

ROSE, SHARON; and RACHEL WALKER. 2004. A Typology of Consonant Agreement by Correspondence. *Language* 80.475–531.

SHIH, STEPHANIE S. to appear. Super additive similarity in Dioula tone harmony. *Proceedings of the West Coast Conference on Formal Linguistics 33*.

SHIH, STEPHANIE S; and SHARON INKELAS. 2014. A subsegmental correspondence approach to contour tone (dis)harmony patterns. *Proceedings of Phonology 2013*. University of Massachusetts, Amherst. http://journals.linguisticsociety.org/proceedings/index.php/amphonology/article/view/22.

SMITH, JENNIFER L. 2011. Category-specific effects. *The Blackwell Companion to Phonology*, ed. by Marc van Oostendorp, Colin Ewen, Elizabeth Hume, and Keren Rice, 2439–2463. Malden, MA: Wiley-Blackwell.

WAYMENT, ADAM. 2009. Assimilation as attraction: Computing distance, similarity, and locality in phonology. Baltimore, MD: Johns Hopkins University ph.d. dissertation.

WILSON, COLIN. 2006. Learning Phonology with Substantive Bias: An Experimental and Computational Study of Velar Palatalization. *Cognitive Science* 30.945–982.

WILSON, COLIN; and MARIEKE OBDEYN. 2009. Simplifying subsidiary theory: statistical evidence from Arabic, Muna, Shona, and Wargamay.

ZHANG, JIE. 2004. Contour tone licensing and contour tone representation. *Language and Linguistics* 5.925–968.

ZHANG, JIE. 2007. Contour tone distribution is not an artifact of tonal melody mapping. *Studies in the Linguistic Sciences* 33.

ZOLL, CHERYL. 2003. Optimal Tone Mapping. *Linguistic Inquiry* 34.225–268.